DOCUMENT RESUME

ED 065 033 HE 003 161

AUTHOR Centra, John A.
TITLE What Instructors Learn from Students: Student and Faculty Ratings of Instruction.
INSTITUTION Educational Testing Service, Princeton, N.J.
PUB DATE 72
NOTE 12p.; Paper presented at the American Educational Research Association Annual Meeting, Chicago, April 1972

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Effective Teaching; *Higher Education; Student Teacher Relationship; *Teacher Evaluation; *Teacher Improvement; *Teaching Quality

ABSTRACT

The purpose of this study was to investigate two questions pertaining to student evaluation of teachers. These are: (1) Do students indeed provide the instructor with information about instructional practices that he doesn't already know? and (2) If this is the case, to what extent is it true at a variety of colleges and for a significant proportion of instructors? The study was conducted by comparing student ratings or descriptions of instruction with the teachers' own self-reported descriptions. The correlation between the 2 sets of descriptions or ratings were not particularly high, indicating only modest agreement in the way faculty and students perceived instruction. In particular, instructors and students did not agree on the extent to which students are free to ask questions or give opinions in class, on the extent to which instructors are concerned with student learning, on the amount of agreement between objectives and what is being taught, on instructor openness to other viewpoints, on the extent to which instructors inform students of how they would be evaluated, on whether the instructor encourages students to think for themselves, and on the clarity of course objectives. These discripancies between the 2 sets of ratings not only underscore the need for student feedback, but also suggest specific areas of instruction where feedback is most essential. (HS)

5.6 N-N.P.O
S.P

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

FILMED FROM BEST AVAILABLE COPY

WHAT INSTRUCTORS LEARN FROM STUDENTS: STUDENT AND FACULTY RATINGS OF INSTRUCTION[1]

John A. Centra

Educational Testing Service

ED 065033

Introduction

The extensive use of student ratings of instruction at the college level over the past several decades has had as its primary purpose the improvement of instruction. Student evaluations typically are seen only by the instructor and are intended to help improve their teaching. Underlying this use are several assumptions, among which is that instructors are learning something new from students about their teaching.

Do students indeed provide the instructor with information about instructional practices that he doesn't already know? And if this is the case, to what extent is it true at a variety of colleges and for a significant proportion of instructors?

The purpose of this study was to investigate these questions by comparing student ratings or descriptions of instruction with the teacher's own self-reported descriptions. Discrepancies between the two sets of ratings, if found, would not only underscore the need for student feedback but would also suggest specific areas of instruction where the feedback is most essestial. In addition, self-ratings have been suggested by some as a worthwhile mode of instructional evaluation. More knowledge about their relationship with student ratings could provide a better understanding of self-ratings, and possibly lead to ways in which they could serve as another source of classroom evaluation.

HE003161

---

[1]This study was supported in part by a grant from the Esso Education Foundation.

Paper read at the American Educational Research Association convention, Chicago, 1972.

Procedure

The sample for the study consisted of 343 teaching faculty at five institutions of higher education. The five institutions included two state colleges, one of which had a predominantly black enrollment, a selective liberal arts college, a multipurpose college, and an urban community college. None of these institutions had, at the time of the study, a systematic program to collect student ratings, nor did a significant portion of their faculty collect student ratings on their own. The majority of teachers in this study, therefore, were not familiar with how students might rate their instruction.

Teachers were randomly selected to participate in the study, and between 75-90% of those selected from each college participated.

A 21-item instructional report questionnaire was used in the study. Included were items that faculty members in an earlier study had identified as providing information they would like to receive from students (Centra, 1972).[2] Among the dimensions of instruction included were the organization of the course, student-teacher interaction, instructor communication, student effort, and stimulation of students. Previous factor analytic studies had identified several of these as dimensions that effectively differentiated among instructors (Isaacson, McKeachie, et al., 1964; Hodgson, 1958; Gibb, 1955; Coffman, 1954).

Responses to 17 of the items were on a four point agree-disagree scale, with a not applicable option also provided. The four remaining items used a four or five point scale with different response options for each item.

---

[2] The form actually contained 23 items, 21 of which could be reworded for reasonable instructor self-ratings.

The wording for each of the statements in the questionnaire differed slightly for students and instructors. For example, an item on course objectives was worded as follows for each group:

Students: The instructor's objectives have been made clear

Teachers: I feel my objectives for the course have been made clear to students

Teachers were asked to "describe this course, your teaching, or the students enrolled." They were told that the reason for obtaining this self-report was to see which items were tapping information already known to most instructors.

The data were collected at mid-semester of the fall 1971 term. Instructors administered the rating form in one class of their own choosing, with the understanding that only they would receive a summary of their students' responses.

Faculty-student comparisons were made primarily in two ways. First, for each item the instructors' self-descriptions were correlated with the mean responses of students in their class (N=343). In this way the relationship between the two sets of ratings could be noted. Second, the differences between the way faculty as a group and students as a group rated or described instruction were investigated by a comparison of means. For this analysis responses for all faculty were compared to the average of the student class means on each item.

Results and Discussion

The results of the comparison of means and the correlational analysis for items 5-21 are presented in Table 1. The correlation between the two sets of descriptions or ratings were not particularly high, indicating only modest agreement in the way faculty and students perceived instruction. While most of

Table 1

Faculty-Student Comparisons to Instructional Report Questionnaire,

343 Classes at Five Colleges[1]

		Mean Faculty Response	Mean of Mean Student Responses	T Test of Means	Number of Colleges Item was significant	Correlations[2]
5	Course objectives made clear	1.55	1.81	7.52*	4	.25
6	Agreement between objectives and teaching	1.51	1.82	9.34*	5	.19
7	Instr. using class time well	1.56	1.72	4.68*	1	.11
8	Instr. availability for students	1.47	1.62	4.32*	2	.28
9	Instr. knows when students don't understand	1.77	1.98	5.43*	3	.21
10	Instr. encourages students to think	1.42	1.71	8.87*	4	.23
11	Instr. concern with student learning	1.26	1.68	13.54*	5	.17
12	Instr. comments helpfully on papers or exams	1.75	2.03	5.74*	2	.33
13	Instr. raises challenging questions	1.68	1.90	5.34*	3	.22
14	Students are free to question or give opinions	1.15	1.67	18.93*	5	.16
15	Instr. preparation for each class	1.47	1.52	1.59	0	.13
16	Instr. informs students of how evaluated	1.52	1.84	8.03*	5	.42
17	Instr. summarizes or emphasizes major points	1.56	1.73	4.34*	2	.13
18	Student interest stimulated by course	1.85	2.01	3.90*	1	.32
19	Students putting effort into course	2.09	1.97	-3.10*	0	.33
20	Instr. openness to other viewpoints	1.42	1.72	8.74*	5	.16
21	Instr. accomplishing objectives for the course	1.70	1.69	- .19	0	.15

*Significant at .01 level.

[1]The N for each item was often less than 343 due to "Not Applicable" instructor responses, i.e., they did not think the item applied to their course.

[2]Correlations between each faculty member's response to each item and the mean response of students in his class. For an N of 343, r of .14 is significant at the .01 level.

the items were statistically significant due to the large N (343), the median correlation was only .21.

Also listed in Table 1 are the mean faculty responses for the items, the mean of the student means, the results of the T tests, and the number of colleges where the difference between the means was significant. Responses for items 5-21 could range from one for "strongly agree", to four for "strongly disagree;" thus, lower values represent greater agreement with each statement. The comparisons of the mean values indicate that instructors as a group generally rated or described their teaching more favorably than did their students. In particular instructors and students did not agree on the extent to which students are free to ask questions or give opinions in class (item 14), on the extent to which instructors are concerned with student learning (11), on the amount of agreement between objectives and what is being taught (6), on instructor openness to other viewpoints (20), on the extent to which instructors inform students of how they would be evaluated (16), on whether the instructor encourages students to think for themselves (10), and on the clarity of course objectives (5). For each of these seven items, instructor-student differences were notable at either four or all five of the colleges.

On the other hand there was little difference between the faculty and student groups in their ratings of the instructor preparation for class (15) and on the extent to which course objectives were being accomplished (21). For the remaining eight items, the differences were modest and in many instances not significant within a college.

But probably more important than a simple comparison of the way an average instructor and an average class rated instruction is some knowledge of how many instructors saw themselves far differently than their students did. A distribution of the differences between each instructor's responses and those of his

class (i.e., the class means) provides that information. Presented in Table 2 is a summary of the results of such a distribution. For each item, the percentage of instructors who gave themselves "considerably poorer" or "considerably better" ratings is indicated within each college and for the total sample. A difference of .63 or greater was used to define "considerably poorer or better" because a difference of at least that great would appear to be large enough to have some practical significance; it is also the approximate standard deviation for most of the student item responses.

For most of the items, between a fourth and a third of the instructors described or rated themselves considerably better than their students did. The median, in fact, was just under 30 percent for all 343 instructors and their classes. Forty-one percent of the instructors gave themselves better ratings on item 14: students are free to ask questions or give opinions in class; and 36 percent on item 11: the instructor is concerned about whether students learn and tries to be actively helpful. Both items deal with faculty-student interaction as do items 8, 9, 10, and 16 for which fairly high percentages of instructors also gave themselves better ratings. The faculty-student interaction dimension, then, appears to be one on which a sizable number of instructors and their students do not agree, and on which student reactions would appear to be especially crucial. Other similar areas would be the instructor's openness to other viewpoints (item 20) and the agreement between announced objectives for the course and what was being taught (6).

A surprisingly large percentage of instructors rated themselves <u>poorer</u> than students did in a few areas. Fifteen percent rated themselves more poorly on class preparation and 12 percent were less satisfied that they were accomplishing course objectives. In general, however, only between four to eight percent of the teachers gave themselves considerably poorer ratings.

Table 2

Results of the Distribution of Differences between Faculty-Student Responses

to the Instructional Report Questionnaire

	Percentage of instructors who gave themselves: Considerably poorer ratings than the mean of students in their class[1]						Considerably better ratings than the mean of students in their class[1]					
	College						College					
	1 N=50	2 N=99	3 N=66	4 N=95	5 N=37	Total N=343	1 N=50	2 N=99	3 N=66	4 N=95	5 N=37	Total N=343
5 Course objectives made clear	6	4	5	2	1	4	28	25	23	31	46	30
6 Agreement between objectives and teaching	2	9	0	2	0	4	44	31	38	36	32	37
7 Instr. using class time well	4	12	5	7	11	8	16	18	15	26	30	21
8 Instr. availability for students	8	9	8	4	30	10	32	24	15	21	22	23
9 Instr. knows when students don't understand	4	8	6	5	14	7	34	21	37	22	28	28
10 Instr. encourages students to think	2	5	5	11	14	4	34	23	14	35	43	29
11 Instr. concern with student learning	2	9	6	1	6	5	36	24	33	45	33	36
12 Instr. comments helpfully on papers or exams	4	8	10	5	7	7	30	28	18	31	41	31
13 Instr. raises challenging questions	4	7	5	10	7	7	28	26	14	23	33	24
14 Students are free to question or give opinions	0	3	0	2	0	2	38	36	42	41	47	41
15 Instr. preparation for each class	6	14	14	19	16	15	20	14	21	20	16	16
16 Instr. informs students of how evaluated	2	10	5	5	0	6	28	28	32	30	42	32
17 Instr. summarizes or emphasizes major points	2	14	7	7	11	9	18	18	33	28	41	28
18 Student interest stimulated by course	8	6	8	7	9	8	20	18	23	16	38	21
19 Students putting effort into course	24	12	19	19	14	18	12	9	8	8	14	10
20 Instr. openness to other viewpoints	4	6	5	3	5	5	30	25	37	33	38	32
21 Instr. accomplishing objectives for the course	4	14	16	9	10	12	10	14	16	16	34	17

[1] A difference of .63 or greater (on a four-point scale) was defined as a "considerably poorer" rating (i.e., tend to disagree with item) or a "considerably better" (i.e., tend to agree with item) rating.

One of the items in the form was unique in that it elicited opinions on student effort in the course (19). For students, the exact wording was: I have been putting a good deal of effort into this course; for instructors it was worded: Students seem to be putting a good deal of effort into this course. The results for this item, as one might expect, were much different than those for other items. Compared to students' responses, 18 percent of the faculty thought students generally were putting considerably less effort into the course, while 10 percent gave students better ratings on effort than students gave themselves. In other words, in this instance students have tended to give themselves better ratings just as instructors did on so many of the previous items.

An inspection of the differences within each college indicates fairly similar results with the exception of college five. In comparison to the other four colleges, higher percentages of the instructors at college five rated themselves considerably better than did their students on a majority of the items. While it is not possible to conclude much on the basis of one college, it is interesting to note that college five was the smallest and most selective of the colleges in the study. Moreover, instructors at college five were given the poorest student ratings among the five colleges, whereas their self-ratings were not much different or poorer than those of instructors elsewhere. Thus the gap between instructor-student ratings at college five was due to the poorer ratings by students, perhaps because of higher expectations on their part, rather than on better ratings by instructors.

Presented in Table 3 is a summary of responses to the first four items, which used varied responses rather than agree-disagree options. The items deal with the pace, the level of difficulty, and the work load of the course, as well

Table 3

Faculty-Student Comparisons at Five Colleges and Total (N = 343),
For Four Items in Instructional Report Questionnaire

	Percentage Responding											
	Students						Faculty					
	College						College					
	1	2	3	4	5	Total	1	2	3	4	5	Total
1 Pace at which material is covered:												
Very or somewhat slow	9	10	7	8	6	9	22	24	10	8	14	16
Very or somwhat fast	26	20	27	23	33	25	20	28	24	30	30	27
2 Level of difficulty of course for students enrolled:												
Very or somewhat elementary	11	13	10	10	9	11	10	7	10	4	8	7
Very or somewhat difficult	31	25	32	21	38	30	26	31	37	37	41	34
3 Work load of course relative to others:												
Lighter	18	22	17	19	18	19	25	24	21	17	14	20
Heavier	20	21	27	29	27	25	35	23	32	32	33	30
4 Extent to which examples and illustrations were used:												
Frequently	60	70	76	67	58	67	88	75	86	82	65	80
Occasionally	28	26	20	26	34	26	12	21	14	18	32	19
Seldom	10	4	4	6	8	6	0	2	0	0	3	1
Never	2	1	1	1	1	1	0	2	0	0	0	1

as the extent to which the instructor used examples and illustrations. Once again there were student-instructor differences although they were not particularly large. Instructors tended to think they more often used examples and illustrations, and at three of the colleges instructors more likely considered the pace at which material was covered to be slow. College five, the selective liberal arts college, was once again noteworthy in that its faculty and to some extent the students reported less frequent use of examples or illustrations in courses.

Summary and Conclusions

A comparison of students' ratings of instruction with teachers' self-reported ratings in over 300 classes at five colleges disclosed a modest relationship between the two set of evaluations. The median correlation for 17 items was .21, indicating that faculty members generally evaluate or describe their teaching quite differently from the way it is evaluated or described by students. Not surprisingly, the highest correlations occurred for the more factual items, on which there was somewhat less chance for disagreement (e.g., the instructor informs students of how they would be evaluated), while items eliciting opinions (e.g., the instructor is using class time well) resulted in the lowest correlations.

Clark and Blackburn (1971) recently reported a similarly low to moderate correlation of .19 between students and faculty self-ratings on a single overall measure of teaching effectiveness. Thus, whether instructors give themselves an overall rating of teaching effectiveness or rate more specific instructional practices, they apparently do not agree with evaluations being made by students. In fact, overall self-ratings and colleague ratings, according to Clark and Blackburn's data, correlated only slightly better (r = .28).

In addition to the general lack of agreement between self and student evaluations, there was also a tendency for teachers as a group to give themselves better ratings than students gave them. This discrepancy, however, was most notable for between a fourth to a third of the 343 instructors in the study, and in particular for items related to student-instructor interaction, course objectives and the instructor's openness to other viewpoints. These areas of instruction, then, would seem to be particular ones in which a sizable proportion of teachers could profit from student feedback.

In conclusion, the results of this study would appear to argue for the collection of student ratings as a means of providing many instructors with information they do not already have about their teaching. Whether instructors actually use this new information is, of course, another question.

References

Centra, J. A. The Student Instructional Report: Its Development and Uses. SIR Report Number 1, Princeton, N. J.: Educational Testing Service, 1972.

Clark, M. J., and Blackburn, R. Assessment of faculty performance: Some correlates between self, colleagues, students, and administrators (1971, submitted for publication).

Coffman, W. E. Determining students' concepts of effective teaching from their ratings of instructors. Journal of Educational Psychology, 1954, 45, 277-285.

Gibb, C. A. Classroom behavior of the college teacher. Educational and Psychological Measurement, 1955, 15, 254-263.

Hodgson, T. F. The general and primary factors in student evaluation of teaching ability. University of Washington, 1958.

Isaacson, McKeachie, et al. The dimensions of student evaluations of teaching. Unpublished mimeo, University of Michigan, 1964.